第27卷 第6期 工 程 数 学 学 报 Vol. 27 No. 6
2010年12月 CHINESE JOURNAL OF ENGINEERING MATHEMATICS Dec. 2010

文章编号:1005-3085(2010)06-1133-04

基于状态转换图的公共场所行人视频监控方法研究

幺居标， 王冬琳

(北京电子科技职业学院，北京 100029)

摘 要: 自动视频监控对于公共场所的安全控制等方面具有重要意义。本文实现了一个实用的行人监控系统，采用了基于改进 Kalman 滤波的背景更新方法进行背景的有效更新，同时开创性地提出了基于状态转换图的行人计数系统，用于处理场景中前景遮挡、粘连等复杂情况，弥补了复杂条件下前景运动物体检测存在的准确度不高的问题，实验结果证明了此方法的有效性。

关键词: 视频监控；Kalman 滤波；状态转换图

分类号: **AMS**(2000) 68T10 **中图分类号**: O23; O29 **文献标识码**: A

1 引言

公共场所的人流量、客流量的统计与测量对于人流的高峰控制告警，提取公共场所的统计数据有很大意义，也对行人通道线路的设计，服务设施地点的选择等问题有着至关重要的指导作用。目前实用的视频监控技术仍然以基于背景模型的前景检测辅以运动物体跟踪为主[1]。最基本的运动物体包括运动分割[2] 与跟踪[3] 两部分。

运动物体之间的粘连与遮挡是非常普遍的现象[4]，尤其对于密集人流监控系统这样的复杂情况。实际问题的特点在于计数要求不是十分精确，人流的向单一集中。我们的工作就是实现一个这样的自动行人计数系统，完成实际应用的要求。本文的其余章节安排如下：首先在第 2 章详细介绍我们提出的基于状态转换图的视频监控系统，第 3 章展示实现的系统的实验结果，最后在第 4 章给出总结。

2 基于状态转换图的公共场所行人视频监控系统

我们实现的系统流程如下：首先通过改进 Kalman 滤波的方法得到实时更新的背景图像，再通过传统的形态学方法得到前景行人；在行人跟踪的问题上，我们通过状态转换图实现计数的多少又是基于自适应面积域值的分析。具体细节在下面详细介绍。

2.1 基于 Kalman 滤波的背景更新与连通区检测

背景提取算法以 Kalman 滤波为基础，Kalman 滤波的普遍公式如下

$$
\begin{aligned}
\hat{B}(k+1) &= \hat{B}(k) + K(k+1)\big[C(k+1) - W\hat{B}(k)\big],\\
K(k+1) &= P(k+1)W^T\big[\partial I + WP(k)W^T\big],\\
P(k+1) &= \big[1 - K(k+1)W\big]P(k)/\partial,
\end{aligned}
\tag{1}
$$

式中 $C(k+1)$ 是 $k+1$ 时刻的输入图像；$\hat{B}(k)$, $\hat{B}(k+1)$ 分别是 k 时刻和 $k+1$ 时刻的估值，也就是背景图像，$K(k+1)$ 是 $k+1$ 时刻的时变增益因子，取值介于 0 和 1 之间；$P(k)$, $P(k+1)$ 分别

收稿日期: 2009-02-23. **作者简介**: 幺居标 (1960年3月生)，男，博士，副教授. 研究方向：智能交通和视频监控.

是 k 时刻和 $k+1$ 时刻的方差矩阵，W 是观测矩阵，∂ 是渐消记忆因子，I 是单位矩阵。若认为图像每一个时空点在空间独立，则上面方程式中的所有变量均退化为标量：$W=1$ 且 $I=1$。同时为了简化计算，假设原递推过程为一个时不变随机过程，P 与 K 均不随 t 的变化改变。联立求解得到 K，上面的公式可以转化为

$$\begin{aligned}\hat{B}(k+1) &= \hat{B}(k) + K(k+1)\big[C(k+1) - \hat{B}(k)\big],\\ K(k+1) &= 1-\partial,\end{aligned} \tag{2}$$

这里的 $K(k+1)$，在系统中称为“更新系数”(update rate)，对这一变量的控制，体现了背景更新的速度和对变化的敏感程度。我们对 Kalman 的传统算法作了改进，加入了非线性的部分。在整个视频的最前几帧，用简单平均的方法得到背景的初步估计，之后如果背景与当前帧差异过大，很大可能上是前景的遮挡，因此不能直接用 Kalman 方法更新背景，为了防止场景中突然出现一个物体并停止转化为前景，如果一个像素点的灰度值一段时间与背景具有同符号的灰度差异，立即更新背景。在一个时间周期内，为了保证背景图案的稳定，$K(k+1)$ 以指数级递减，而经过一个周期，为了防止背景的突变造成的影响，又将 $K(k+1)$ 的值恢复至初始状态。这一过程可以用下式表示

$$K(k+1) = \mathrm{e}^{-\beta(t \text{ MOD } T)}. \tag{3}$$

2.2 状态转换图的实现

在新帧到来后，算法将根据新的连通域的情况，更新跟踪信息，跟踪信息通过状态转换图来管理：每一个跟踪信息对应一个状态，从当前帧分析的结果，决定跟踪信息的下一个状态，使用状态转换图可以有效处理前景连通区的合并、分离等各种情况。状态转换图如图 1 所示。

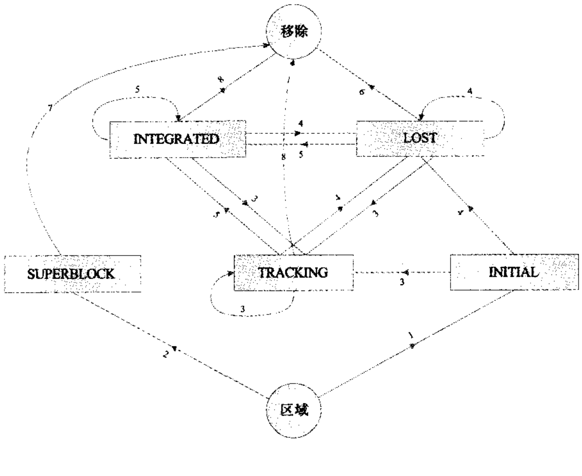

图 1: 状态转换图

图中箭头上的数字代表状态转换的各个条件，将在下面详述。首先将其中各个状态的意义说明如下：

INITIAL：正常初始化，除了 SUPERBLOCK，这是区块进入跟踪信息的唯一途径；

TRACKING：单一的正在被跟踪的区块；

LOST：跟踪区块丢失，系统允许跟踪区块丢失若干帧，再从丢失的状态中恢复；

INTEGRATED：两个或多个区块在跟踪中合成一个大区块的情况，这时多个区块都被标记为 INTEGRATED；

SUPERBLOCK：超级块，当视频上的行人过于密集，会形成前景区域非常大的情况，这种情况区域产生区与销毁区都不再起作用，这时将前景区域作为SUPERBLOCK处理。

下面分别介绍各个状态转换的条件：

条件1是前景区域进入跟踪链的入口，这个条件在所有原始的跟踪信息处理过之后判断。系统定义产生区与销毁区都是场景周围的环形区域，当一个当前区域在所有的跟踪处理过后仍未被跟踪且其中心在生成区内，条件1成立，前景区域进入跟踪链且被标记为INITIAL。条件2由区域延伸的范围与区域的面积判断，条件3判断并跟踪单一区块。算法将在预测方向上查找新的跟踪区域的中心。如果在预测的位置附近最近的区块中心到预测位置的距离在一个阈值以内则进行跟踪。当在预测位置处本身没有区域存在时，条件4成立，认为这一区块跟踪暂时丢失。当两个行人相互靠近但在预测位置上存在有效前景时，条件5成立，认为有两个区块合并发生。条件6判断何时认为暂时丢失的区域“完全丢失”，即从跟踪链中移除。这里通过连续丢失的总帧数判断，如果连续丢失的总帧数超过某一域值，则丢弃这一跟踪的区域信息。条件7是移除超级块的条件：如果当前所有的连通域中不再有满足“超级块”条件的区域，则将“超级块”移除。条件8是计数的关键，当监测到一个有效的跟踪链(TRACKING或INTEGRATED)到达图像的销毁区，发消息计数并移除跟踪区域信息。

在进行计数时，需要确定区域所包含的行人个数。首先考虑两人的简单情况，当两个人重叠遮挡时，一般总面积会小于两个单人面积，因此有必要引入一个系数，用于标记视野中由于两个目标物体相互遮挡造成的前景面积损失。我们采用如下的模型来进行模拟：假设相互遮挡的目标物体具有相同的大小(设为单位1)，并且处于同一直线上，建立模型：在长度为2的固定线段上投下两条长度为1的线段，同时假设投放的线段位置均匀分布且互相独立，则两条投放的线段在长度为2线段上遮住长度的数学期望为

$$E(x) = 1 + E\big(|a-b|\big) = \frac{4}{3}, \tag{4}$$

因此，两个行人重叠时平均每人占据面积的最佳估计值为0.667 (2/3) 倍单人占据面积。

3 实验结果

背景提取的对比实验结果如图2所示，其中图2左图展示了直接使用Kalman滤波的背景图效果，在图像中存在“鬼影”的现象，而相应的提出的方法可以较好的去除这一现象。

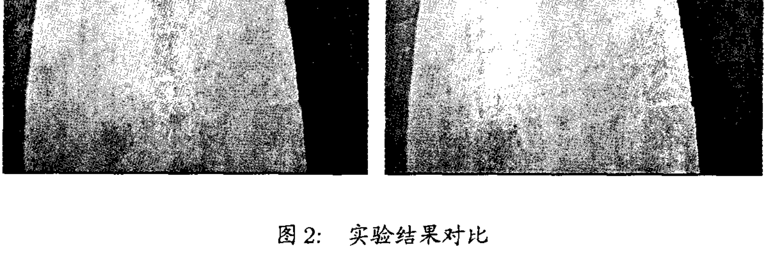

图2: 实验结果对比

同时，我们对设计的系统进行了实际在线测试，使用的测试数据视频长度为7分43秒，经过总人数为下行732，分辨率320 × 240，帧率29.97 fps。通过对测试数据的多次重复实验，我们得到了如表1的统计实验结果。其中，“下行”栏代表系统实际检测到的下行人流数目，

“其他”栏显示的是由于噪声引起的系统检测到的其他方向的人流数目。5次实验的统计数据列在了表格的右方。

表1: 实验数据测试结果

实验次数	1	2	3	4	5	手工	平均值	标准差	均值误差	最大误差
下行	759	759	717	763	712	732	742	22.56	1.37%	4.23%
其他	7	10	17	10	12	0	11.2	3.31	–	–

从实验结果可以看出，本文的方法有效地达到了行人流量视频监测的目的，在地铁站这种人流比较密集的地区，误检率和漏检率相对较少，可以达到人流监测、预警的实用目的。

4 讨论与结论

本文致力于在行人流量密集的公共场所采用传统的前景/背景方式进行有效的视频监控，并采用了状态转换图与设定状态之间转换规则的方式进行运动物体的跟踪，来解决目标之间的遮挡与分开的问题。需要强调的是，状态转换图方法具有较好的可扩展性，可以灵活应用于各种不同场景与条件下的视频监控情况，我们可以设计其他的状态与状态转移条件来满足实际情况的需要。

参考文献:

[1] Collins R T, *et al.* Introduction to the special section on video surveillance[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(7): 745-746

[2] Elgammal A M, Davis L S. Probabilistic framework for segmenting people under occlusion[C]// Proceedings of Eighth IEEE International Conference on Computer Vision, Vol II, 2001: 145-152

[3] Zhao T, Nevatia R. Tracking multiple humans in complex situations[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004, 26(9): 1208-1221

[4] Conte D, *et al.* A graph-based, multi-resolution algorithm for tracking objects in presence of occlusions[J]. Pattern Recognition, 2006, 39(4): 562-572

Pedestrian Video Surveillance Approach Based on the State-transition Graph

YAO Ju-biao, WANG Dong-lin

(Beijing Vocational College of Electronic Science, Beijing 100029)

Abstract: The automatic visual surveillance plays an important role in the security control of public places. In this paper, a practical pedestrian surveillance system is proposed, the system is based on the improved Kalman filter background update approach. We also implement a new state-transition graph system to solve some complicate situations, like foreground occlusion and conjunction. Experimental results show the efficiency of our approaches.

Keywords: video surveillance; Kalman filter; state-transition graph

Received: 23 Feb 2009. **Accepted:** 24 Aug 2009.